VARIANCE COMPONENTS ESTIMATION: A THUMBNAIL REVIEW

by

BU-612-M May, 1977

S. R. Searle

Cornell University, Ithaca, New York

Abstract

This paper has two distinct parts. The first is a brief account of early work (1939-1953) on variance component estimation and of some recent uses and applications of variance components models. The second part is a summary account, in note form, of methods currently available for estimating variance components, particularly from unbalanced data (having unequal numbers of observations in the subclasses).

PART I. HISTORY, AND CURRENT USES

I.1. Variance components models

There is widespread familiarity with the traditional analysis of variance model such as that for the completely randomized design:

$$y_{ij} = \mu + \alpha_i + e_{ij} \ . \tag{1}$$

In this equation y_{ij} is the j^{th} observation on the i^{th} treatment, with μ representing a general mean and α_i the effect of the i^{th} "treatment". The expected value of y_{ij} is taken as $E(y_{ij}) = \mu + \alpha_i$ for E representing expectation over repeated sampling. The term e_{ij} in (1) represents the difference $y_{ij} - E(y_{ij})$ and is usually taken as being a random variable (usually called residual, or error, or both), with

Paper No. BU-612-M in the Biometrics Unit Mimeo Series, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14853.

zero mean and variance σ_e^2 for all $i = 1, \cdots, a$ and $j = 1, \cdots, n$, the individual e_{ij}'s also being uncorrelated.

The parameters of interest in this model are the mean μ and the α_i's, the effects of the "treatments" on the yield; and one object of analyzing the data is that of estimating linear functions of μ and the α_i's. Best linear unbiased estimators of two functions of interest are, for example,

$$\widehat{\mu + \alpha_i} = \bar{y}_{i.} = \sum_{j=1}^{n} y_{ij}/n \, .$$

and

$$\widehat{\alpha_i - \alpha_{i'}} = \bar{y}_{i.} - \bar{y}_{i'.} \, .$$

In the context of this model the μ and the α_i's are taken as being constants, albeit unknown and unknowable, but nevertheless fixed constants. As such, they are usually called <u>fixed effects</u>. They are deemed to be constants representing the effects of the different "treatments" being studied. The "treatments" are the things of particular interest, chosen by some investigator because of his interest in them: different diets fed to laboratory animals, farm livestock or to humans, different fertilizers given to a corn crop, different forage crops grown in the same region, different machines used in a manufacturing process, different drugs given for the same illness, and so on. The possibilities are legion — as are the varieties of models and their complexities, reaching far beyond those of (1).

Now consider the α_i's of equation (1) as being realized (but unobservable) values of a random variable having zero mean and variance σ_α^2, with the α_i's being uncorrelated with each other and with the e_{ij}'s. [In this case $E(y_{ij}) = \mu$, and e_{ij} is defined as $e_{ij} = y_{ij} - E(y_{ij}|\alpha_i)$ where $E(y_{ij}|\alpha_i)$ is a conditional expected value.] In this context the individual α_i's are no longer things of particular interest as they are in fixed effects models; those that occur in the data are deemed to be just a random sample of α's selected from a population defined as having zero mean and variance σ^2_α. There is therefore little or no reason for estimating either the α_i's or differences between them; the parameter of interest so far as they are concerned is now σ^2_α. Because in this case (1) gives $\sigma^2_y = \sigma^2_\alpha + \sigma^2_e$, the variances σ^2_α and σ^2_e, being components of the variance of y, are called variance components. This use of (1) leads to the model being called a variance components model, and the α_i's are called random effects. Correspondingly, the model is sometimes called the random model.

Some models have both fixed effects and random effects, in which case the name mixed model is used. An example would be a randomized complete block design having model equation

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

where the α_i's are fixed effects representing treatments and the β_j's are random effects representing blocks. The parameters of interest would be μ and the α_i's and the variance components σ^2_β and σ^2_e.

I.2. History, 1939-1953

The basic principle for estimating variance components has been, and to a large extent still is, that of equating quadratic functions of the observations to their expected values. Obvious candidates for such functions are those of the analysis of variance table. The first papers describing this procedure appear to be those of Daniels [1939], whose interest was in weights of slubbings from the carding process in the woolen industry, and Winsor and Clarke [1940] who analyzed catches of different species in successive hauls of plankton nets. The papers were published only a few months apart and it seems certain they were the results of independent work by the respective authors. Both papers give expected sums of squares for two or three different analyses of variance, Winsor and Clarke explicitly showing

$$E \sum_{i=1}^{a} n(\bar{y}_{i.} - \bar{y}_{..})^2 = (a - 1)(n\sigma_\alpha^2 + \sigma_e^2)$$

and (2)

$$E \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2 = a(n - 1)\sigma_e^2$$

for the random model based on (1). They then estimate σ_α^2 and σ_e^2 by $\hat{\sigma}_e^2$ and $\hat{\sigma}_\alpha^2$, the solutions to the equations

$$\sum_{i=1}^{a} n(\bar{y}_{i.} - \bar{y}_{..})^2 = (a - 1)(n\hat{\sigma}_\alpha^2 + \hat{\sigma}_e^2)$$

and (3)

$$\sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2 = a(n - 1)\hat{\sigma}_e^2 .$$

Daniels does not mention R. A. Fisher in regard to deriving expected values of sums of squares, whereas Winsor and Clarke describe their derivation of (2) as being "a straightforward extension of the suggestions of R. A. Fisher in his 'Statistical Methods for Research Workers' [Sec.] 40." Presumably this is the Seventh Edition, published in 1938, wherein Sec. 40 is the section dealing with the intraclass correlation, exactly as does the same section, unchanged, in the Twelfth Edition of 1954. The important suggestion of Fisher's is in table 39 which, although he makes no explicit mention of expectation whatever, contains exactly the results (2). Nor does he give any serious attention to estimating σ_α^2 beyond saying "we may make estimates of the values A and B [σ_α^2 and σ_e^2], or in other words we may analyze the variance into the portions contributed by the two causes". Both Daniels and Winsor and Clarke use the expectation notation and are concerned with estimating σ_α^2 and σ_e^2.

At about the same time as the Daniels and Winsor and Clarke papers were published (the latter in what, even at that time, must have been somewhat of an obscure journal for statisticians), Snedecor's third edition [1940] became available with, as far as I can see, no reference to variance components at all. Page 205 contains discussion of estimating the intra-class correlation as A/(A + B), just as does Fisher [1938]. The nearest thing to characterizing A as a variance component is the description that "A is the same for all ··· samples – it is the common element, analogous to covariance." And that is, of course, the case: the covariance between y_{ij} and $y_{ij'}$ for $j \neq j'$ is σ^2_α.

Winsor and Clarke not only use (2) and (3), for balanced data, but they also derive the expectations

$$E \sum_{i=1}^{a} n_i(\bar{y}_{i.} - \bar{y}_{..})^2 = (n_. - \sum_{i=1}^{a} n_i^2/n_.)\sigma^2_\alpha + (a-1)\sigma^2_e$$

and (4)

$$E \sum_{i=1}^{a} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 = (n_. - a)\sigma^2_e \ ,$$

for unbalanced data, something which Daniels [1939] does not address himself to. Interestingly enough, Snedecor [Third Edition, 1940] touches obliquely on this subject in Example 10.21 (p. 205) where, in referring to unbalanced data of Table 10.8, he asks the question "Why can't you calculate intra-class correlation accurately" for such data? Winsor and Clarke's results (4) would show that you could. Needless to say, that example does not appear in the sixth edition, Snedecor and Cochran [1967].

Although Daniels [1939] and Winsor and Clarke [1940] represent both sides of the Atlantic, it appears that major developments in variance components estimation subsequently took place mainly in the U. S. A. An exception to this was Ganguli

[1941], dealing with nested classifications, and then came Crump [1946] concerned with randomized complete blocks, and Satterthwaite [1946] dealing with approximate sampling distributions of estimated variance components. This was followed by Eisenhart [1947] who put a firm foundation to the distinction between the fixed effects model (Model I) and the random effects model (Model II), a distinction which Yates [1967] later took great exception to. Sampling variances of estimators obtainable from (4) were given in Hammersley [1949] for arbitrary distributional properties, and in a doctoral thesis by Crump [1947] for normality assumptions. These results from the thesis were included in Crump's [1951] review paper, but not those concerned with maximum likelihood estimation, a topic which reasserted itself in Hartley and Rao [1967] and has been actively pursued ever since, with no immediate end in sight (e.g., Harville [1977]).

Anderson and Bancroft [1952] is the first book that gives any treatment of variance components, its final four chapters being devoted to the topic entirely. This really set the subject on a firm footing and well and truly laid out the procedure of equating analysis of variance sums of squares to their expectations as a method of estimating variance components. The book deals very thoroughly with estimation from unbalanced data for both mixed and random models; it also deals with unbalanced data for nested classifications and, after considering incomplete blocks designs, it poses a number of pertinent research problems, many of which have still not been answered satisfactorily. In all, the book is a milestone in variance components estimation.

Active interest in estimation from balanced data continued well into the 1960's, by which time several optimum properties of the estimators had been established (see references in Searle [1971], for example, particularly those by Graybill and co-workers). Estimation from unbalanced data in crossed classifications, and mixtures of crossed and nested classifications, with mixed or random models, got its prime start from Henderson [1953]. Active interest in unbalanced data has continued unabated to the present day and has not subsided yet.

I.3. Uses and applications

Most statistical methods are developed in response to the demands of practical problems. Variance components estimation is no exception. The first papers, by Daniels [1939] and Winsor and Clarke [1940], dealt with woolen industry and with plankton net data, respectively. Crump [1946] was interested in Drosophila egg production and he also refers to a variety of other applications of variance components: enumeration sampling, cereal experiments, swine breeding (three papers), corn breeding and soil sampling. Papers by Hazel and Terrill [e.g., 1945] on sheep breeding could be added to the list. Clearly, by the mid-40's, animal and plant breeders were making considerable use of variance components. The Anderson and Bancroft [1952] book also contains references to numerous uses of variance components in subject-matter disciplines: industrial experimentation, corn trials, psychological testing, sample surveys (wheat fields, soybean trials and forest nurseries), the sampling of baled wool, studies of egg production and hog prices, and analyses of the efficacy of measuring instruments.

Thus we see that, like many statistical techniques founded in practical problems, the early literature of variance components estimation contains plentiful reference to those problems. In contrast, today's statistical literature deals very largely with just the mathematics of statistical methodology, with much less space being devoted to practical problems and accompanying data than was the case twenty and more years ago. This is certainly true of variance component estimation no less than it is of other topics in statistics. Nevertheless, literature of the subject-matter disciplines (and occasionally of statistics) continues to bear witness to the uses for variance components estimates. In addition to the frequent and established uses in genetics and animal breeding such as estimating genetic parameters (e.g., Becker et al. [1977]) and using them in prediction (Searle [1974]) there are now uses in a variety of other disciplines. Closely allied to the animal breeder's parameter of repeatability is the psychologist's and educationalist's measure of reliability of a test instrument, namely $\sigma^2_{\text{respondent}}/(\sigma^2_{\text{respondent}} + \sigma^2_{\text{residual}})$, as, for example, in Alwin [1976]. Geneticists and others who use the experimental design of the diallel cross (originating in genetics) also make great use of variance components – Randall [1976] provides a comprehensive review – and so do those designing sample surveys. Analyses of trajectory and orbital data in rocket flight testing have been based on the mixed model version of variance components models (e.g., Bush [1971]) and so have analyses of data from clinical trials involving several clinics (Chakravorti and Grizzle [1975]). Kalman filtering techniques of engineering, as described by Duncan and Horn [1972], also utilize mixed model theory, as noted by Harville [1977]. And economists nowadays make very wide use of mixed models in combining cross-section with time series data (e.g., Houthakker et al. [1974]), referring to their models as error components models. Variance components estimation continues, therefore, to be a technique that is quite widely used in data analysis, as well as receiving considerable attention on its theoretical side.

I.4. References

Alwin, D. F. [1976]. Attitude scales as cogeneric tests: a re-examination of an attitude-behavior model. Sociometry 39, 377-383.

Anderson, R. L. and Bancroft, T. A. [1952]. Statistical Theory in Research, 1st Ed., McGraw Hill.

Becker, W. A., Spencer, J. V., Verstrate, J. A., and Minosh, L. M. [1977]. Genetic analysis of chicken egg yolk cholesterol. Poultry Science 56, 895-901.

Bush, N. [1971]. Unmodeled error analysis on trajectory and orbital estimation. Technometrics 13, 303-314.

Chakravorti, S. R. and Grizzle, J. E. [1975]. Analysis of data from multiclinic experiments. Biometrics 31, 325-338.

Crump, S. L. [1946]. The estimation of variance components in analysis of variance. Biometrics Bulletin 2, 7-11.

Crump, S. L. [1947]. The estimation of components of variance in multiple classifications. Ph. D. Thesis, Iowa State University, Ames, Iowa.

Crump, S. L. [1951]. Present status of variance components analysis. Biometrics 7, 1-16.

Daniels, H. E. [1939]. The estimation of components of variance. J. Roy. Stat. Soc. Supp. 6, 186-197.

Duncan, D. B. and Horn, S. D. [1972]. Linear dynamic recursive estimation from the viewpoint of regression analysis. J. Am. Stat. Assoc. 67, 815-821.

Eisenhart, C. [1947]. The assumptions underlying the analysis of variance. Biometrics 3, 1-21.

Fisher, R. A. [1938, 1954]. Statistical Methods for Research Workers, 7th and 12th Editions, Oliver and Boyd, Edinburgh.

Ganguli, M. [1941]. A note on nested sampling. Sankhyā 5, 449-452.

Hammersley, J. M. [1949]. The unbiased estimate and standard error of the interclass variance. Metron 15, 189-205.

Hartley, H. O. and Rao, J. N. K. [1967]. Maximum likelihood estimation for the mixed analysis of variance model. Biometrika 54, 93-108.

Harville, D. A. [1977]. Maximum likelihood approaches to variance component estimation and to related problems. J. Am. Stat. Assoc. 72 (in press).

Hazel, L. N. and Terrill, C. E. [1945]. Heritability of weaning weight and staple length in range Rambouillet lambs. J. Animal Sci. 4, 347-358.

Henderson, C. R. [1953]. Estimation of variance and covariance components. Biometrics 9, 226-252.

Houthakker, H. S., Verleger, P. K., Jr., and Sheehan, D. P. [1974]. Dynamic demand analyses for gasoline and residential electricity. American J. of Agric. Econ. 56, 412-418.

Randall, J. J. [1976]. The diallel cross. Unpublished M.S. Thesis, Biometrics Unit, Cornell University, Ithaca, New York.

Satterthwaite, F. E. [1946]. An approximate distribution of estimates of variance components. Biometrics Bulletin 2, 110-114.

Searle, S. R. [1971]. Linear Models, Wiley, New York.

Searle, S. R. [1974]. Prediction, mixed models, and variance components. Reliability and Biometry, Eds. F. Proschan and R. J. Serfling, Soc. of Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 229-266.

Snedecor, G. W. [1940]. Statistical Methods, 3rd Ed., The Collegiate Press, Inc., Ames, Iowa.

Snedecor, G. W. and Cochran, W. G. [1967]. Statistical Methods, 6th Ed., Iowa State University Press, Ames, Iowa.

Winsor, C. P. and Clarke, G. L. [1940]. A statistical study of variation in the catch of plankton nets. Sears Foundation J. Marine Res. 3, 1-34.

Yates, F. [1967]. A fresh look at the basic principles of the design and analysis of experiments. Proc. 5th Berkeley Symp. Math. Statist. Prob. IV, 777-790. L. Lecam and J. Neyman, Eds., University of California Press.

PART II. ESTIMATING VARIANCE COMPONENTS FROM UNBALANCED DATA IN MIXED MODELS OF THE ANALYSIS OF VARIANCE

A summary of methods – in note form.

II.0. Introduction

Confine attention to the 2-way crossed classification model:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$$

$$i = 1 \cdots a \quad j = 1 \cdots b \quad k = 1 \cdots n_{ij} \quad \sum_{ij}\sum n_{ij} \equiv N$$

Fixed effects model

Balanced data

All $n_{ij} = n$: the analysis of variance is familiar.

Mean	1	$abn\bar{y}^2_{...}$
Rows	a-1	$SSA = \Sigma bn\bar{y}^2_{i..} - abn\bar{y}^2_{...}$
Columns	b-1	$SSB = \Sigma an\bar{y}^2_{.j.} - abn\bar{y}^2_{...}$
Interaction	(a-1)(b-1)	$SSAB = \Sigma\Sigma n\bar{y}^2_{ij.} - \Sigma bn\bar{y}^2_{i..} - \Sigma an\bar{y}^2_{.j.} + abn\bar{y}^2_{...}$
Residual	ab(n-1)	$SSE = \Sigma\Sigma\Sigma y^2_{ijk} - \Sigma\Sigma n\bar{y}^2_{ij.}$
Total	abn	$\Sigma\Sigma\Sigma y^2_{ijk}$

Unbalanced data: s cells containing data.

2 partitionings of sums of squares.

Rows before columns		OR	Columns before rows	
$R(\mu)$	1		$R(\mu)$	1
$R(\alpha\|\mu)$	a-1		$R(\beta\|\mu)$	b-1
$R(\beta\|\mu,\alpha)$	b-1		$R(\alpha\|\mu,\beta)$	a-1
$R(\gamma\|\mu,\alpha,\beta)$	s-a-b+1		$R(\gamma\|\mu,\alpha,\beta)$	s-a-b+1
SSE	N-s		SSE	N-s
Total	N		Total	N

Mixed model:

β_j's remain as fixed effects

α_i's random

$E(\alpha_i) = 0$

$var(\underset{-}{\alpha}) = \sigma^2_\alpha I_a$

γ_{ij}'s random

$E(\gamma_{ij}) = 0$

$var(\underset{-}{\gamma}) = \sigma^2_\gamma I_s$

s = ab for balanced data

Want to estimate: μ, β's, σ^2_α, σ^2_γ and σ^2_e

Balanced data:

Use part of analysis of variance table for fixed effects model

$$E(SSA) = (a-1)(bn\sigma^2_\alpha + n\sigma^2_\gamma + \sigma^2_e)$$

$$E(SSAB) = (a-1)(b-1)(n\sigma^2_\gamma + \sigma^2_e)$$

$$E(SSE) = ab(n-1)\sigma^2_e$$

Estimators

$$SSA = (a-1)(bn\hat{\sigma}^2_\alpha + n\hat{\sigma}^2_\gamma + \hat{\sigma}^2_e)$$

$$SSAB = (a-1)(b-1)(n\hat{\sigma}^2_\gamma + \hat{\sigma}^2_e)$$

$$SSE = ab(n-1)\hat{\sigma}^2_e$$

Properties of estimators: unbiased

minimum variance quadratic unbiased

under normality, minimum variance unbiased

Unbalanced data:

Variety of methods available, several based on same principle as preceding:

Develop $\underset{\sim}{q}$ as a vector of quadratic forms in $\underset{\sim}{y}$.

Derive $E(\underset{\sim}{q})$; each element will be a linear combination of variance components, elements of $\underset{\sim}{\sigma}^2$.

$$E(\underset{\sim}{q}) = \underset{\sim}{C}\underset{\sim}{\sigma}^2 \text{ for some } \underset{\sim}{C}$$

Estimation: $\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{C}^{-1}\underset{\sim}{q}$

Question: What quadratics are used as elements for $\underset{\sim}{q}$?

II.1. Analysis of variance method (Henderson's [1953] Method 1)

This method uses quadratic forms analogous to sums of squares of balanced data ANOVA, e.g.,

$$SSA^* = \sum_i n_{i.}\bar{y}^2_{i..} - N\bar{y}^2_{...}$$

$$SSAB^* = \sum\sum_{ij} n_{ij}\bar{y}^2_{ij.} - \sum_i n_{i.}\bar{y}^2_{i..} - \sum_j n_{.j}\bar{y}^2_{.j.} + N\bar{y}^2_{...}$$

Note: $SSAB^*$ is not positive definite; it is not a sum of squares.

Estimation: equate SS^*'s to expectations

Properties: easy to compute

unbiased for random models

sampling variances available for 1, 2 and 3-way classifications

not unbiased for mixed models, because the fixed effects, β_j's, occur in $E(SS^*)$'s.

History: Henderson [1953]: described method

Searle [1971a]: collected details for 1-, 2- and 3-way classifications, including sampling variances of estimates.

II.2. Henderson's [1953] Method 2

Designed to overcome biasedness of Method 1 for mixed models.

Retains relative ease of computing.

Principle: "Correct" data for fixed effects.

Use Method 1 on corrected data.

Make slight adjustments.

$$\underset{\sim}{y} = \underbrace{\underset{\sim\sim}{Xb}}_{\text{fixed}} + \underbrace{\underset{\sim\sim}{Zu}}_{\text{random}} + \underset{\sim}{e}$$

Use normal equations as if $\underset{\sim}{u}$ were fixed:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix} \begin{bmatrix} b^{o} \\ u^{o} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Correct for $\underset{\sim}{b}$: use a $\underset{\sim}{b}^{o}$ such that

$$z = y - Xb^{o} = \mu^{*}1 + Zu + Ke, \text{ for some } K.$$

Use Method 1 on $\underset{\sim}{z}$ just as if it were $\underset{\sim}{y}$ without fixed effects.

Adjustments: To coefficients of σ_e^2 in E(SS's), to account for $\underset{\sim}{K}$.

Condition: No interactions, fixed-by-random.

History: Henderson [1953]: first described, but not clearly.

Searle [1968]: generalized, clarified, decried as not invariant.

Henderson, Searle and Schaeffer [1974]: established invariance, and described computing procedure.

II.3. Fitting constants method (Henderson's [1953] Method 3)

Use R( )'s of fitting constants for fixed effects models

$$E\ R(\alpha,\gamma|\mu,\beta) = c_1\sigma_\alpha^2 + c_1\sigma_\gamma^2 + (s-b)\sigma_e^2$$

$$E\ R(\gamma|\mu,\alpha,\beta) = c_2\sigma_\gamma^2 + (s-a-b+1)\sigma_e^2$$

$$E\ SSE = (N-s)\sigma_e^2$$

or, if no interaction

$$E\ R(\alpha|\mu,\beta) = c_1\sigma_\alpha^2 + (a-1)\sigma_e^2$$

$$E\ SSE = (N-a-b+1)\sigma_e^2$$

Properties: Unbiased.

Reduces to ANOVA for balanced data.

Difficulties: Can be difficult to compute (i.e., inverting large matrices).

Not uniquely defined: can have more equations than variance components, e.g., for 2-way random model, can use

$R(\alpha\|\mu)$		$R(\beta\|\mu)$		$R(\beta\|\mu,\alpha)$
$R(\beta\|\mu,\alpha)$	OR	$R(\alpha\|\mu,\beta)$	OR	$R(\alpha\|\mu,\beta)$
$R(\gamma\|\mu,\alpha,\beta)$		$R(\gamma\|\mu,\alpha,\beta)$		$R(\gamma\|\mu,\alpha,\beta)$
SSE		SSE		SSE
$y'y - N\bar{y}^2$		$y'y - N\bar{y}^2$		

History: Henderson [1953]: described method.

Rohde and Tallis [1969]: give general expressions for sampling variances.

II.4. Henderson's Mixed Model Equations (MME)

A general form of the mixed model is

$$\underset{\sim}{y} = \underbrace{\underset{\sim}{X}\underset{\sim}{b}}_{\text{fixed}} + \underbrace{\underset{\sim}{Z}\underset{\sim}{u}}_{\text{random}} + \underset{\sim}{e} \quad \text{with} \quad E(\underset{\sim}{u}) = \underset{\sim}{0} \quad \operatorname{var}(\underset{\sim}{u}) = \underset{\sim}{D}$$

$$E(\underset{\sim}{e}) = \underset{\sim}{0} \quad \operatorname{var}(\underset{\sim}{e}) = \underset{\sim}{R}$$

$$\text{and} \quad E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{b} \quad \operatorname{var}(\underset{\sim}{y}) \equiv \underset{\sim}{V} = \underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{Z}' + \underset{\sim}{R}$$

GLS for $\underset{\sim}{b}$:

$$\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}\underset{\sim}{b}^{\circ} = \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} \tag{1}$$

Difficulty: $\underset{\sim}{V}^{-1}$ has order N.

Assuming $\underset{\sim}{u}$ fixed, GLS for $\underset{\sim}{b}$ and $\underset{\sim}{u}$:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z \end{bmatrix} \begin{bmatrix} \tilde{b} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \tag{2}$$

Amend equations by adding $\underset{\sim}{D}^{-1}$ to $\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}$:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ X'R^{-1}Z & Z'R^{-1}Z + D^{-1} \end{bmatrix} \begin{bmatrix} b^* \\ u^* \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \tag{3}$$

These are Henderson's mixed model equations (MME).

Advantage: The MME $\underset{\sim}{b}^*$ of (3) requires less computing than the GLS $\underset{\sim}{b}^{\circ}$ of (1); and $\underset{\sim}{b}^* \equiv \underset{\sim}{b}^{\circ}$.

History: Henderson *et al.* [1959]: described, showed $\underset{\sim}{b}^* = \underset{\sim}{b}^{\circ}$, and discussed $\underset{\sim}{u}^*$.
Lindley and Smith [1972]: equations arise in Bayesian setting.
Henderson [1973]: MME related to MINQUE and ML.
Harville [1976]: renewed attention.

Special case: A single random factor: $\underset{\sim}{u} = \underset{\sim}{\alpha}$, $\underset{\sim}{D} = \sigma_\alpha^2 \underset{\sim}{I}_a$ and $\underset{\sim}{R} = \sigma_e^2 \underset{\sim}{I}_N$.

Define $\lambda \equiv \sigma_e^2/\sigma_\alpha^2$ and $\underset{\sim}{P} \equiv \underset{\sim}{Z}'\underset{\sim}{Z} + \lambda\underset{\sim}{I}$. Then
$$\begin{bmatrix} X'X & X'Z \\ Z'X & P \end{bmatrix} \begin{bmatrix} b^* \\ u^* \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}. \tag{4}$$

II.5. Thompson's iterative method

Models with only 1 random factor; e.g., 2-way classification without interaction

$$y_{ijk} = \underbrace{\mu + \beta_j} + \alpha_i + e_{ijk}$$
$$\underset{\sim}{y} = \underset{\sim\sim}{Xb} + \underset{\sim\sim}{Z\alpha} + \underset{\sim}{e}$$

Fitting constants method: Based on

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix} \begin{bmatrix} b^o \\ \alpha^o \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad \text{and} \quad R(\mu,\alpha,\beta) = (b^{o\prime} \quad \alpha^{o\prime}) \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

$$\hat{\sigma}_e^2 = \frac{y'y - R(\mu,\alpha,\beta)}{N-a-b+1} \qquad \hat{\sigma}_\alpha^2 = \frac{R(\alpha|\mu,\beta) - (a-1)\sigma_e^2}{N - \Sigma\Sigma n_{ij}^2/n_{i.}}$$

Cunningham and Henderson [1968]: Used mixed model equations

$$\begin{bmatrix} X'X & X'Z \\ Z'X & P \end{bmatrix} \begin{bmatrix} b^* \\ \alpha^* \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad \text{and} \quad R^*(\mu,\alpha,\beta) = (b^{*\prime} \quad u^{*\prime}) \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

and got

$$\overset{*}{\sigma}{}_e^2 = \frac{y'y - R^*(\mu,\alpha,\beta)}{N-a-b+1} \qquad \overset{*}{\sigma}{}_\alpha^2 = \frac{R^*(\mu,\alpha,\beta) - R(\mu,\beta) - (a-1)\overset{*}{\sigma}{}_e^2}{N + a\lambda - \Sigma\Sigma n_{ij}^2/n_{i.}}$$

Iterate on $\lambda = \sigma_e^2/\sigma_\alpha^2$ with $\underset{\sim}{P} \equiv \underset{\sim}{Z}'\underset{\sim}{Z} + \lambda\underset{\sim}{I}$.

Thompson's [1969] method:

Located error in expectations of Cunningham and Henderson; correction yields

$$\tilde{\sigma}_e^2 = \frac{y'y - R^*(\mu,\alpha,\beta)}{N - b} \qquad \tilde{\sigma}_\alpha^2 = \frac{R^*(\mu,\alpha,\beta) - R(\mu,\beta)}{N - \Sigma\Sigma n_{ij}^2/n_{i.}}$$

Iterate on $\lambda = \sigma_e^2/\sigma_\alpha^2$.

Computing formulae for 2-way, no interaction: Searle [1973].

Extension to 2-way, with interaction: Corbeil and Searle [1973].

(This is an extension from 1 to 2 random factors.)

II.6. MIVQUE and MINQUE (Four papers by C. R. Rao)

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \sum_{\theta=A}^{K+1} \underset{\sim}{Z}_\theta \underset{\sim}{u}_\theta \qquad \theta = \text{factors } A, B, \cdots, K, \text{ including interactions}$$

$\underset{\sim}{u}_{K+1} \equiv \underset{\sim}{e}$ and $\underset{\sim}{Z}_{K+1} \equiv \underset{\sim}{I}_N$; $E(\underset{\sim}{u}_\theta) = \underset{\sim}{0}$; $\text{var}(\underset{\sim}{u}_\theta) = \sigma^2_\theta \underset{\sim}{I}_{N_\theta}$; and $\text{cov}(\underset{\sim}{u}_\theta \underset{\sim}{u}'_\varphi) = \underset{\sim}{0}$ for $\theta \neq \varphi$.

Notation: $\underset{\sim}{V}_\theta = \underset{\sim}{Z}_\theta \underset{\sim}{Z}'_\theta$; $\quad \underset{\sim}{V}_+ = \sum_{\theta=A}^{K+1} \underset{\sim}{V}_\theta$, (Rao's $\underset{\sim}{V}$); $\quad \underset{\sim}{V} = \sum_{\theta=A}^{K+1} \sigma^2_\theta \underset{\sim}{V}_\theta$, (Rao's $\underset{\sim}{V}^*$).

Estimation: by quadratics $\underset{\sim}{y}'\underset{\sim}{A}\underset{\sim}{y}$ with $\underset{\sim}{A}\underset{\sim}{X} = \underset{\sim}{0}$ for invariance to $\underset{\sim}{b}$, choosing $\underset{\sim}{A}$:

MIVQUE: to minimize $2\text{tr}(VA)^2$ + term in $\underset{\sim}{A}$ and kurtosis parameters [1971a, p. 268; 1972, p. 113].

Under normality, minimize $\text{tr}(VA)^2$: [1971b, pp. 447, 453].

$$\underset{\sim}{R} = \underset{\sim}{V}^{-1} - \underset{\sim}{V}^{-1}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1}$$

$\underset{\sim}{S} = \{s_{\theta\varphi}\} = \{\text{tr}(\underset{\sim}{V}_\theta \underset{\sim}{R} \underset{\sim}{V}_\varphi \underset{\sim}{R})\}$ and $\underset{\sim}{t} = \{t_\theta\} = \{\underset{\sim}{y}'\underset{\sim}{R}\underset{\sim}{V}_\theta \underset{\sim}{R}'\underset{\sim}{y}\}$ for $\theta, \varphi = A, B, \cdots, K+1$.

Solve

$\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{S}^{-1}\underset{\sim}{t}$, with $\underset{\sim}{\sigma}^{2\prime} = (\sigma^2_A \; \sigma^2_B \cdots \sigma^2_K \; \sigma^2_e)$, involved in $\underset{\sim}{V}$, and hence in $\underset{\sim}{R}$, $\underset{\sim}{S}$ and $\underset{\sim}{t}$.

MIVQUE proper (normality): for pre-assigned $\underset{\sim}{\sigma}^2_0$, $\hat{\underset{\sim}{\sigma}}^2 \equiv \hat{\underset{\sim}{\sigma}}^2(\underset{\sim}{\sigma}^2_0)$ is locally MIVQUE for $\underset{\sim}{\sigma}^2_0$. $\underset{\sim}{\sigma}^2_0 = \underset{\sim}{1}$ is often suggested, $\underset{\sim}{V}$ then being $\underset{\sim}{V}_+$.

Iterative MIVQUE (normality): iterate $\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{S}^{-1}\underset{\sim}{t}$ on $\underset{\sim}{\sigma}^2$.

MINQUE: minimize a Euclidean norm (without normality).
MINQUE = MIVQUE (normality).

History: Townsend [1968] and Townsend and Searle [1971]: did initial work.
Harville [1969]: dropped zero mean assumption.
Rao [1970, 1971a, 1971b, 1972]: developed MIVQUE and MINQUE.
LaMotte [1973]: made generalizations.
Maddala and Mount [1974]: comparative numerical study.
Harville [1975, 1977]: summarizes and comments.

II.7. Maximum likelihood

Use normality and same model as MIVQUE:

$$\underset{\sim}{y} = \underset{\sim\sim}{Xb} + \sum_{\theta=A}^{K} \underset{\sim}{Z}_\theta \underset{\sim}{u}_\theta + \underset{\sim}{e} .$$

Notation:

$$\gamma_\theta = \sigma_\theta^2/\sigma_e^2 \qquad\qquad \underset{\sim}{H} = \underset{\sim}{I}_N + \sum_{\theta=A}^{K} \gamma_\theta \underset{\sim}{Z}_\theta \underset{\sim}{Z}_\theta'$$

$$\text{var}(\underset{\sim}{y}) = \sigma_e^2 \underset{\sim}{H} = \underset{\sim}{V} .$$

Equations:

$$\underset{\sim}{X}'\underset{\sim}{\tilde{H}}^{-1}\underset{\sim\sim}{X\tilde{b}} = \underset{\sim}{X}'\underset{\sim}{\tilde{H}}^{-1}\underset{\sim}{y}$$

$$\tilde{\sigma}_e^2 = (\underset{\sim}{y} - \underset{\sim\sim}{X\tilde{b}})'\underset{\sim}{\tilde{H}}^{-1}(\underset{\sim}{y} - \underset{\sim\sim}{X\tilde{b}})/N$$

$$\text{tr}(\underset{\sim}{\tilde{H}}^{-1}\underset{\sim}{Z}_\theta \underset{\sim}{Z}_\theta') = (\underset{\sim}{y} - \underset{\sim\sim}{X\tilde{b}})'\underset{\sim}{\tilde{H}}^{-1}\underset{\sim}{Z}_\theta \underset{\sim}{Z}_\theta' \underset{\sim}{\tilde{H}}^{-1}(\underset{\sim}{y} - \underset{\sim\sim}{X\tilde{b}})/\tilde{\sigma}_e^2, \quad \text{for } \theta = A, \cdots, K .$$

For unbalanced data these equations have no solution; neither do they for some balanced data situations (e.g., 2-way crossed classification, random model, with interaction). Solutions must be confined to positive values.

History: Hartley and Rao [1967]: established equations; numerical solution by steepest descent.

Hartley and Vaughn [1972]: computer program, and small examples.

Hemmerle and Hartley [1973]: Newton-Raphson, and the W-transformation.

Miller [1973]: improved iterative procedure.

Harville [1975, 1977]: a comprehensive review.

Hemmerle and Lorens [1976]: improved the W-transformation.

II.8. REML: Restricted Maximum Likelihood

Use normality and model: $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \sum_{\theta=A}^{K} \underset{\sim}{Z}_{\theta}\underset{\sim}{u}_{\theta} + \underset{\sim}{e}$.

Method: Maximize likelihood of a transformation of $\underset{\sim}{y}$. Use a transformation that lacks $\underset{\sim}{b}$ in its model.

Most general description:

Use $\underset{\sim}{w} = \underset{\sim}{A}\underset{\sim}{y}$, for $\underset{\sim}{A}$ being any full row rank matrix of order $[N - r(X)] \times N$ with $\underset{\sim}{A}\underset{\sim}{X} = \underset{\sim}{0}$.

Elements of $\underset{\sim}{w}$ are called error contrasts: $E(\underset{\sim}{w}) = \underset{\sim}{0}$.

Can use any $N - r$ linearly independent error contrasts: their log likelihoods differ only by additive constants.

$$L = -\tfrac{1}{2}\log|\underset{\sim}{V}| - \tfrac{1}{2}\log|\underset{\sim}{X}^{*\prime}\underset{\sim}{V}^{-1}\underset{\sim}{X}^{*}| - \tfrac{1}{2}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{b}^{\circ})'\underset{\sim}{V}^{-1}(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{b}^{\circ}),$$

for $b^{\circ} = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y}$

and $\underset{\sim}{X}^{*}$ being any r linearly independent columns of $\underset{\sim}{X}$.

Special cases: $\underset{\sim}{A}$ such that $\underset{\sim}{A}\underset{\sim}{A}' = \underset{\sim}{I}$ and $\underset{\sim}{A}'\underset{\sim}{A} = \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'$;

$\underset{\sim}{A}$ such that $\underset{\sim}{A}'(\underset{\sim}{A}\underset{\sim}{A}')^{-1}\underset{\sim}{A} = \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'$.

History: Patterson and Thompson [1971]: unbalanced data, largely b.i.b. designs.

Harville [1974]: shows results for any error contrasts.

Hocking and Kutner [1975]: simulate results on b.i.b. designs.

Harville [1975, 1977]: gives L, and makes comprehensive review.

Corbeil and Searle [1976a]: use a particular $\underset{\sim}{A}$ ($\underset{\sim}{T}$ in their notation).

Corbeil and Searle [1976b]: give analytic comparisons of REML and ML for balanced data, and numeric comparisons for unbalanced data.

II.9. Dispersion-Mean Correspondence

Model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \Sigma \mathbf{Z}_i \mathbf{u}_i$$

$$E(\mathbf{y}) = \mathbf{X}\mathbf{b}$$

$$\text{var}(\mathbf{y}) = \mathbf{V} = \Sigma \mathbf{Z}_i \mathbf{Z}_i' \sigma_i^2 = \Sigma \mathbf{V}_i \sigma_i^2, \quad \mathbf{V}_i = \mathbf{Z}_i \mathbf{Z}_i'$$

$$\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^- \mathbf{X}', = \mathbf{M}' = \mathbf{M}^2 \text{ and } \mathbf{M}\mathbf{X} = \mathbf{0}$$

$$\mathbf{D} = [\text{vec}(\mathbf{V}_1) \ \ \text{vec}(\mathbf{V}_2) \cdots \text{vec}(\mathbf{V}_k)] .$$

It can be shown that

$$E(\mathbf{M}\mathbf{y} * \mathbf{M}\mathbf{y}) = (\mathbf{M} * \mathbf{M})\mathbf{D}\boldsymbol{\sigma}^2 .$$

Because $\mathbf{M}\mathbf{y} * \mathbf{M}\mathbf{y}$, a Kronecker product of vectors, is a vector, this is effectively a linear model that provides opportunity for estimating $\boldsymbol{\sigma}^2$.

OLS $\hat{\sigma}^2$:

$$\hat{\boldsymbol{\sigma}}^2 = (\mathbf{D}_M'\mathbf{D}_M)^- \mathbf{D}_M'(\mathbf{y} * \mathbf{y}) \quad \text{for} \quad \mathbf{D}_M \equiv (\mathbf{M} * \mathbf{M})\mathbf{D} .$$

GLS $\tilde{\sigma}^2$:

$$\tilde{\boldsymbol{\sigma}}^2 = (\mathbf{D}'\mathbf{F}_M^-\mathbf{D})^- \mathbf{D}'\mathbf{F}_M^-(\mathbf{y} * \mathbf{y})$$

for

$$\mathbf{F}_M = (\mathbf{M} * \mathbf{M})\mathbf{F}(\mathbf{M} * \mathbf{M})$$

and

$$\mathbf{F} = \text{var}[(\mathbf{y} - \mathbf{X}\mathbf{b}) * (\mathbf{y} - \mathbf{X}\mathbf{b})] .$$

History: Pukelsheim [1976]: developed this approach.

II.10. Relationships among Methods

(1) ANOVA ≡ Henderson 1 (Definition).

(2) ML estimators are ML solutions <u>subject</u> to non-negativity conditions.

Balanced Data

(3) ANOVA = Henderson 2 = Henderson 3 = REML = MINQUE (MIVQUE).

(4) Some ML equations have no closed form solution.

When solutions do exist, some (but not all) = ANOVA.

(Differences occur in "degrees of freedom").

Unbalanced Data

ML and MINQUE (MIVQUE) model:

$$\underset{\sim}{y}_{N\times 1} = \underset{\sim}{X}\underset{\sim}{b} + \sum_{\theta=A}^{K} \underset{\sim}{Z}_\theta \underset{\sim}{u}_\theta + \underset{\sim}{e}, \quad \text{with } \underset{\sim}{u}_\theta \text{ order } n_\theta \times 1.$$

Henderson's mixed model equations (MME's):

$$\gamma_\theta = \sigma^2_\theta/\sigma^2_e \qquad \underset{\sim}{D} = \text{diag}\{\gamma_\theta \underset{\sim}{I}_{n_\theta}\} \qquad \theta = A, \cdots, K \qquad \underset{\sim}{R} = \sigma^2_e \underset{\sim}{I}_N$$

$$\begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{X} & \underset{\sim}{X}'\underset{\sim}{Z} \\ \underset{\sim}{Z}'\underset{\sim}{X} & \underset{\sim}{Z}'\underset{\sim}{Z} + \underset{\sim}{D}^{-1} \end{bmatrix} \begin{bmatrix} \underset{\sim}{b}^* \\ \underset{\sim}{u}^* \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{y} \\ \underset{\sim}{Z}'\underset{\sim}{y} \end{bmatrix}$$

with solution defined as

$$\begin{bmatrix} \underset{\sim}{b}^* \\ \underset{\sim}{u}^* \end{bmatrix} = \begin{bmatrix} \underset{\sim}{C}_{00} & \underset{\sim}{C}_{01} \\ \underset{\sim}{C}_{10} & \underset{\sim}{C}_{11} \end{bmatrix} \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{y} \\ \underset{\sim}{Z}'\underset{\sim}{y} \end{bmatrix}$$

where

$$\underset{\sim}{C}_{11} = \{\underset{\sim}{C}_{\theta\phi}\}; \quad \underset{\sim}{u}^* = \{\underset{\sim}{u}_\theta\}; \quad \theta,\phi = A, \cdots, K .$$

MINQUE (MIVQUE) and the MME's

MINQUE equations (MIVQUE under normality)

$$\underset{\sim}{S}\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{t}; \text{ i.e., } \{s_{\theta\phi}\}\hat{\underset{\sim}{\sigma}}^2 = \{\underset{\sim}{t}_\theta\} \text{ for } \theta,\phi = A, \cdots, K, e$$

are given by

(5) $$s_{\theta\theta} = [tr(\underset{\sim}{C}_{\theta\theta})/\gamma_\theta^2 - 2tr(\underset{\sim}{C}_{\theta\theta})/\gamma_\theta + n_\theta]/\gamma_\theta^2$$

$$s_{\theta\phi} = tr(\underset{\sim}{C}_{\theta\phi}\underset{\sim}{C}_{\phi\theta})/\gamma_\theta^2\gamma_\phi^2$$

$$s_{\theta e} = s_{e\theta} = tr[\underset{\sim}{C}_{\theta\theta} - \sum_{\phi=A}^{K} tr(\underset{\sim}{C}_{\theta\phi}\underset{\sim}{C}_{\phi\theta})/\gamma_\phi]/\gamma_\theta$$

$$s_{ee} = N - r(\text{MME's}) + \sum_{\theta=A}^{K}\sum_{\phi=A}^{K} tr(\underset{\sim}{C}_{\theta\phi}\underset{\sim}{C}_{\phi\theta})/\gamma_\theta\gamma_\phi$$

$$t_\theta = u_\theta^{*\prime}u_\theta^{*}/\gamma_\theta^2$$

$$t_e = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{b}^{*\prime}\underset{\sim}{X}'\underset{\sim}{y} - \underset{\sim}{u}^{*\prime}\underset{\sim}{Z}'\underset{\sim}{y} - \sum_{\theta=A}^{K}\gamma_\theta q_\theta$$

ML and the MME's

(6) Iterate ML using

$$\tilde{\sigma}_e^2 = \frac{\underset{\sim}{y}'(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{b}^* - \underset{\sim}{Z}\underset{\sim}{u}^*)}{N} \quad \text{and} \quad \tilde{\sigma}_\theta^2 = \frac{\underset{\sim}{u}_\theta^{*\prime}\underset{\sim}{u}_\theta^{*}}{n_\theta - tr(\underset{\sim}{T}_{\theta\theta})}$$

$$\text{for } (\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{Z}\underset{\sim}{D})^{-1} = \{\underset{\sim}{T}_{\theta\phi}\}, \quad \theta,\phi = A, \cdots, K.$$

This iteration always gives positive estimators.

REML and MINQUE (MIVQUE)

(7) REML equations = MINQUE equations.

(8) REML estimators = Iterative MINQUE estimators.

(9) First iterate of REML = A MINQUE estimator.

History:

Patterson and Thompson [1971]: first indication of result for t_θ, for b.i.b. design.
Henderson [1973]: extended results, MME's, ML and MINQUE.
La Motte [1973]: indicated results for REML and MINQUE.
Schaeffer [1975]: published some details of MME's and MINQUE, with many misprints.
Harville [1975, 1977]: comprehensive review.

REFERENCES

The following references are in chronological order within each topic.

1. Analysis of variance method (Henderson's Method 1) – for random, but not mixed, models.

Henderson, C. R. [1953]. Estimation of variance and covariance components. Biometrics 9, 226-252.

Searle, S. R. [1968]. Another look at Henderson's methods of estimating variance components. Biometrics 24, 749-788.

Searle, S. R. [1971a]. Linear Models, John Wiley and Sons, New York.

Searle, S. R. [1971b]. Topics in variance components estimation. Biometrics 27, 1-77.

2. Henderson's Method 2 (for models with no fixed-by-random interactions).

Henderson, C. R. [1953].

Searle, S. R. [1968, 1971a, 1971b].

Henderson, C. R., Searle, S. R., and Schaeffer, L. R. [1974]. The invariance and calculation of Method 2 for estimating variance components. Biometrics 30, 583-588.

3. Fitting constants method (Henderson's Method 3).

Henderson, C. R. [1953].

Searle, S. R. [1968, 1971a, 1971b].

Rohde, C. A. and Tallis, G. M. [1969]. Exact first-order and second-order moments of estimates of components of covariance. Biometrika 56, 517-526.

4. Henderson's Mixed Model Equations.

Henderson, C. R., Kempthorne, O., Searle, S. R., and von Krosigk [1959]. Estimation of environmental and genetic trends from records subject to culling. Biometrics 15, 192-218.

Lindley, D. V. and Smith, A. F. M. [1972]. Bayes estimates for the linear model. J. Roy. Statist. Soc. (B) 34, 1-18.

Henderson, C. R. [1973]. (i) Maximum likelihood estimation of variance components; and (ii) MINQUE of variance components. Unpublished manuscripts, Animal Science Dept., Cornell University.

Harville, D. A. [1976]. Extension of the Gauss-Markov theorem to include the estimation of random effects. Annals of Statist. 4, 384-395.

5. Thompson's iterative method.

Cunningham, E. P. and Henderson, C. R. [1968]. An iterative procedure for estimating fixed effects and variance components in mixed model situations. Biometrics 24, 13-25. Correction 25, 777-778.

Thompson, R. [1969]. Iterative estimation of variance components for non-orthogonal data. Biometrics 25, 767-773.

Searle, S. R. [1971a].

Searle, S. R. [1973]. Computing procedures for estimating variance components from unbalanced data in the 2-way crossed classification, no interaction, mixed model. Paper No. BU-450-M in the Biometrics Unit Mimeo Series, Cornell University.

Corbeil, R. R. and Searle, S. R. [1973]. Iterative estimation of variance components in the 2-way crossed classification mixed model, with interaction, using unbalanced data. Paper No. BU-460-M in the Biometrics Unit Mimeo Series, Cornell University.

6. MIVQUE and MINQUE.

Townsend, E. C. [1968]. Unbiased estimators of variance components in simple unbalanced designs. Ph.D. Thesis, Cornell University, Ithaca, New York.

Harville, D. A. [1969]. Quadratic unbiased estimation of variance components for the one-way classification. Biometrika 56, 313-326. Correction, Biometrika 57, 226, 1970.

Rao, C. R. [1970]. Estimation of heteroscedastic variances in linear models. J. Am. Stat. Assoc. 65, 161-172.

Rao, C. R. [1971a]. Estimation of variance and covariance components — MINQUE theory. J. Multivar. Anal. 1, 257-275.

Townsend, E. C. and Searle, S. R. [1971]. Best quadratic unbiased estimation of variance components from unbalanced data in the 1-way classification. Biometrics 27, 643-657.

Rao, C. R. [1971b]. Minimum variance quadratic unbiased estimation of variance components. J. Multivar. Anal. 1, 445-456.

Rao, C. R. [1972]. Estimation of variance and covariance components in linear models. J. Am. Stat. Assoc. 67, 112-115.

La Motte, L. R. [1973]. Quadratic estimation of variance components. Biometrics 29, 311-330.

Maddala, G. S. and Mount, T. D. [1973]. Comparative study of alternative estimators for variance components models. J.A.S.A. 68, 324-328.

Harville, D. A. [1975, 1977]. Maximum likelihood approaches to variance component estimation and to related problems. Tech. Rpt. #75-0175, Aerospace Research Laboratory, Wright-Patterson Air Force Base, Ohio, 1975. Revised and up-dated, J. Am. Stat. Assoc. 72 (in press), 1977.

7. Maximum Likelihood.

Hartley, H. O. and Rao, J. N. K. [1967]. Maximum likelihood estimation for the mixed analysis of variance model. Biometrika 54, 93-108.

Hartley, H. O. and Vaughn, W. K. [1972]. A computer program for the mixed analysis of variance model based on maximum likelihood. In Statistical Papers in Honor of George W. Snedecor, Ed., T. A. Bancroft, 129-144, Iowa State University Press, Ames.

Miller, J. J. [1973]. Asymptotic properties and computation of maximum likelihood estimates in the mixed model of the analysis of variance. Tech. Rpt. #12 (NR-042-034), Department of Statistics, Stanford University, Stanford, California.

Hemmerle, W. J. and Hartley, H. O. [1973]. Computing maximum likelihood estimates for the mixed A.O.V. model using the W-transformation. Technometrics 15, 819-832.

Hemmerle, W. J. and Lorens, J. A. [1976]. Improved algorithm for the W-transform in variance component estimation. Technometrics 18, 207-212.

Jennrich, R. I. and Sampson, P. F. [1976]. Newton-Raphson and related algorithms for maximum likelihood estimation of variance components. Technometrics 18, 11-18.

8. Restricted Maximum Likelihood.

Patterson, H. D. and Thompson, R. [1971]. Recovery of inter-block information when block sizes are unequal. Biometrika 58, 545-554.

Harville, D. A. [1974]. Bayesian inference for variance components using only error contrasts. Biometrika 61, 383-385.

Hocking, R. R. and Kutner, M. H. [1975]. Some analytical and numerical comparisons of estimators in the mixed A.O.V. model. Biometrics 31, 19-28.

Harville, D. A. [1975, 1977].

Corbeil, R. R. and Searle, S. R. [1976a]. Restricted maximum likelihood (REML) estimation of variance components in the mixed model. Technometrics 18, 31-38.

Corbeil, R. R. and Searle, S. R. [1976b]. A comparison of variance components estimators. Biometrics 32, 779-791.

9. Dispersion-mean correspondence.

Pukelsheim, Friedrich [1976]. Estimating variance components in linear models. J. Multivar. Anal. 6, 626-629.

10. Relationships among ML, REML and MINQUE.

Henderson, C. R. [1973].

La Motte, L. R. [1973].

Harville, D. A. [1975, 1977].

Schaeffer, L. R. [1975]. Disconnectedness and variance component estimation. Biometrics 31, 969-977.